

CLAIMS

What is claimed is:

Sub
A1 ✓
2

A method for re-distributing, over a cluster of one or more active nodes, management of locks on shared resources, the method comprising:

establishing a first master node as master for one or more resources in response to a hash value range being mapped to said first master node, wherein the hash value range is associated with said one or more resources by a hash function; transferring responsibility for mastering said one or more resources from the first master node to a second master node during a transfer time interval; and processing lock requests received at a receiving node of the first master node and the second master node by the receiving node during the transfer time interval.

2. The method of Claim 1, said transferring further comprising:
re-mapping the hash value range to the second master node at the first master node;
sending initial lock information resident on the first master node at a start of the transfer time interval to the second master node;
receiving acknowledgments at the first master node from other active nodes in the cluster; said acknowledgements indicating that said other active nodes have been informed that said second master node is assuming responsibility for mastering said one or more resources.

3. The method of Claim 2, said transferring further comprising, in response to receiving acknowledgements from all active nodes in the cluster by a full acknowledgement

3 time, sending updated lock information resident on the first master node at said full
4 acknowledgment time to the second master node, wherein the transfer time interval
5 ends at an update time of said sending updated lock information.

1 4. The method of Claim 1, said transferring further comprising:
2 receiving initial lock information at the second master node, said initial lock
3 information resident on the first master node at a start of the transfer time
4 interval;
5 re-mapping the hash value range to the second master node at the second master node;
6 and
7 sending a broadcast message to all other nodes in the cluster that the second master
8 node is a new master node for resources associated with the hash value range.

1 5. The method of Claim 1, said transferring further comprising:
2 receiving a broadcast message at a set of nodes in the cluster, the set of nodes
3 including all nodes in the cluster except the first master node and the second
4 master node;
5 wherein the broadcast message indicates that the second master node is a new master
6 node for resources associated with the hash value range;
7 re-mapping the hash value range to the second master node at each node in said set of
8 nodes in the cluster;
9 sending an acknowledgment to the first master node from each node in said set of
10 nodes in response to the broadcast message, said acknowledgement indicating

11 that said each node in said set of nodes has been informed that said second
12 master node is assuming responsibility for mastering said one or more
13 resources; and
14 after sending the acknowledgement, said each node in said set of nodes sending
15 subsequent lock requests for resources associated with the hash value range to
16 the second master node.

1 6. The method of Claim 4, said transferring further comprising receiving updated lock
2 information from the first master node at the second master node wherein the transfer
3 time interval ends at an update time of said receiving updated lock information.

1 7. The method of Claim 6, wherein:
2 lock requests include a sequence number; and
3 said method further comprises deleting stale requests among the updated lock
4 information received at the second master node, the stale requests indicated by
5 sequence numbers earlier than sequence numbers in lock requests already
6 processed on the second master node.

1 8. A computer-readable medium bearing instructions for re-distributing, over a cluster of
2 one or more active nodes, management of locks on shared resources, the instructions
3 arranged to cause one or more processors to perform:

4 establishing a first master node as master for one or more resources in response to a
5 hash value range being mapped to said first master node, wherein the hash
6 value range is associated with said one or more resources by a hash function;
7 transferring responsibility for mastering said one or more resources from the first
8 master node to a second master node during a transfer time interval; and
9 processing lock requests received at a receiving node of the first master node and the
10 second master by the receiving node during the transfer time interval.

1 9. The computer-readable medium of Claim 8, said transferring further comprising:
2 re-mapping the hash value range to the second master node at the first master node;
3 sending initial lock information resident on the first master node at a start of the
4 transfer time interval to the second master node;
5 receiving acknowledgments at the first master node from other active nodes in the
6 cluster; said acknowledgements indicating that said other active nodes have
7 been informed that said second master node is assuming responsibility for
8 mastering said one or more resources.

1 10. The computer-readable medium of Claim 9, said transferring further comprising, in
2 response to receiving acknowledgements from all active nodes in the cluster by a full
3 acknowledgement time, sending updated lock information resident on the first master
4 node at said full acknowledgment time to the second master node, wherein the transfer
5 time interval ends at an update time of said sending updated lock information.

1 11. The computer-readable medium of Claim 8, said transferring further comprising:
2 receiving initial lock information at the second master node, said initial lock
3 information resident on the first master node at a start of the transfer time
4 interval;
5 re-mapping the hash value range to the second master node at the second master node;
6 and
7 sending a broadcast message to all other nodes in the cluster that the second master
8 node is a new master node for resources associated with the hash value range.

1 12. The computer-readable medium of Claim 8, said transferring further comprising:
2 receiving a broadcast message at a set of nodes in the cluster, the set of nodes
3 including all nodes in the cluster except the first master node and the second
4 master node;
5 wherein the broadcast message indicates that the second master node is a new master
6 node for resources associated with the hash value range;
7 re-mapping the hash value range to the second master node at each node in said set of
8 nodes in the cluster;
9 sending an acknowledgment to the first master node from each node in said set of
10 nodes in response to the broadcast message, said acknowledgement indicating
11 that said each node in said set of nodes has been informed that said second
12 master node is assuming responsibility for mastering said one or more
13 resources; and

14 after sending the acknowledgement, said each node in said set of nodes sending
15 subsequent lock requests for resources associated with the hash value range to
16 the second master node.

1 13. The computer-readable medium of Claim 11, said transferring further comprising
2 receiving updated lock information from the first master node at the second master
3 node wherein the transfer time interval ends at an update time of said receiving
4 updated lock information.

1 14. The computer-readable medium of Claim 13, wherein:
2 lock requests include a sequence number; and
3 said instructions causing the one or more processors to further perform deleting stale
4 requests among the updated lock information received at the second master
5 node, the stale requests indicated by sequence numbers earlier than sequence
6 numbers in lock requests already processed on the second master node.

1 15. A computer-readable medium bearing instructions for re-distributing, over a cluster of
2 one or more active nodes, management of locks on shared resources from a first
3 master node to a second master node during a transfer time interval, the instructions
4 arranged to cause one or more processors on the first master node to perform:
5 re-mapping a hash value range initially assigned to the first master node to the second
6 master node, wherein the hash value range is associated with one or more of
7 the shared resources by a hash function;

8 sending initial lock information resident on the first master node at a start of the
9 transfer time interval to the second master node;
10 receiving acknowledgments at the first master node from other active nodes in the
11 cluster; said acknowledgements indicating that said other active nodes have
12 been informed that said second master node is assuming responsibility for
13 mastering said one or more resources.

1 16. The computer-readable medium of Claim 15, said instructions further causing the one
2 or more processors to perform, in response to receiving acknowledgements from all
3 active nodes in the cluster by a full acknowledgement time, sending updated lock
4 information resident on the first master node at said full acknowledgment time to the
5 second master node, wherein the transfer time interval ends at an update time of said
6 sending updated lock information.

1 17. The computer-readable medium of Claim 15, said instructions further causing the one
2 or more processors to perform processing lock requests received during the transfer
3 time interval until receiving acknowledgements from all active nodes in the cluster.

1 18. A computer-readable medium bearing instructions for re-distributing, over a cluster of
2 one or more active nodes, management of locks on shared resources from a first
3 master node to a second master node during a transfer time interval, the instructions
4 arranged to cause one or more processors on the second master node to perform:

5 receiving initial lock information resident on the first master node at a start of the
6 transfer time interval;
7 re-mapping a hash value range initially assigned to the first master node to the second
8 master node, wherein the hash value range is associated with one or more of
9 the shared resources by a hash function; and
10 sending a broadcast message to all other nodes in the cluster that the second master
11 node is a new master node for resources associated with the hash value range.

1 19. The computer-readable medium of Claim 18, said instructions further causing the one
2 or more processors to perform receiving updated lock information from the first
3 master node wherein the transfer time interval ends at an update time of said receiving
4 updated lock information.

1 20. The computer-readable medium of Claim 19, wherein:
2 lock requests include a sequence number; and
3 said instructions further cause the one or more processors to perform deleting stale
4 requests among the updated lock information received at the second master
5 node, the stale requests indicated by sequence numbers earlier than sequence
6 numbers in lock requests already processed on the second master node.

1 21. A computer-readable medium bearing instructions for re-distributing, over a cluster of
2 one or more active nodes, management of locks on shared resources from a first

3 master node to a second master node during a transfer time interval, the instructions
4 arranged to cause one or more processors on a third node to perform:
5 receiving a broadcast message indicating that the second master node is a new master
6 node for resources associated with a hash value range, wherein the hash value
7 range is associated with one or more of the shared resources by a hash
8 function;
9 re-mapping the hash value range to the second master node;
10 sending an acknowledgment to the first master node in response to the broadcast
11 message, said acknowledgement indicating that said third node has been
12 informed that said second master node is assuming responsibility for mastering
13 said one or more resources; and
14 after said sending an acknowledgment, sending subsequent lock requests for the one
15 or more of the shared resources to the second master node.

add
A2